Title
Identification of platform-independent gene expression markers of cisplatin nephrotoxicity.

Journal
Environmental health perspectives, 112(4)

ISSN
0091-6765

Authors
Thompson, Karol L
Afshari, Cynthia A
Amin, Rupesh P
et al.

Publication Date
2004-03-01

DOI
10.1289/ehp.6676

Peer reviewed

Identification of Platform-Independent Gene Expression Markers of Cisplatin Nephrotoxicity

Karol L. Thompson,[1] Cynthia A. Afshari,[2,3] Rupesh P. Amin,[3] Timothy A. Bertram,[4] Bruce Car,[5] Michael Cunningham,[3] Clive Kind,[6] Jeffrey A. Kramer,[4] Michael Lawton,[4] Michael Mirsky,[4] Jorge M. Naciff,[7] Victor Oreffo,[6] P. Scott Pine,[1] and Frank D. Sistare[1]

[1]Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, Maryland, USA; [2]Amgen, Inc., Thousand Oaks, California, USA; [3]National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, North Carolina, USA; [4]Pfizer Inc, St. Louis, Missouri, USA, and Groton, Connecticut, USA; [5]Bristol-Myers Squibb, Wilmington, Delaware, USA; [6]AstraZeneca Research and Development, Charnwood, Leicestershire, United Kingdom; [7]The Procter & Gamble Company, Cincinnati, Ohio, USA

Within the International Life Sciences Institute Committee on Genomics, a working group was formed to focus on the application of microarray technology to preclinical assessments of drug-induced nephrotoxicity. As part of this effort, Sprague-Dawley rats were treated with the nephrotoxicant cisplatin at doses of 0.3–5 mg/kg over a 4- to 144-hr time course. RNA prepared from these animals was run on a variety of microarray formats at multiple sites. A set of 93 differentially expressed genes associated with cisplatin-induced renal injury was identified on the National Institute of Environmental Health Sciences (NIEHS) custom cDNA microarray platform using quadruplicate measurements of pooled animal RNA. The reproducibility of this profile of statistically significant gene changes on other platforms, in pooled and individual animal replicate samples, and in an independent study was investigated. A good correlation in response between platforms was found among the 48 genes in the NIEHS data set that could be matched to probes on the Affymetrix RGU34A array by UniGene identifier or sequence alignment. Similar results were obtained with genes that could be linked between the NIEHS and Incyte or PHASE-1 arrays. The degree of renal damage induced by cisplatin in individual animals was commensurate with the number of differentially expressed genes in this data set. These results suggest that gene profiles linked to specific types of tissue injury or mechanisms of toxicity and identified in well-performed replicated microarray experiments may be extrapolatable across platform technologies, laboratories, and in-life studies. *Key words:* **cisplatin, cross-platform, kidney, microarrays, nephrotoxicity.** *Environ Health Perspect* **112:488–494 (2004). doi:10.1289/txg.6676 available via *http://dx.doi.org/* [Online 15 January 2004]**

The International Life Science Institute (ILSI) Health and Environmental Sciences Institute (HESI) Committee on Application of Genomics and Proteomics to Mechanism-Based Risk Assessment was established to advance the scientific basis for the development and application of genomic and proteomic methodologies to mechanism-based risk assessment (Robinson et al. 2003). One of the long-term objectives of the subcommittee is to relate changes in gene and protein expression to other measures of toxicity. Three toxicity working groups were formed to assess the application of microarray technology to nephrotoxicity, hepatotoxicity, and genotoxicity. The HESI Nephrotoxicity Working Group conducted studies using three nephrotoxicants—cisplatin, gentamicin, and puromycin—with different mechanisms or target regions within the kidney (Kramer et al. 2004). RNA was prepared at one site and sent to other laboratories for analysis on multiple microarray formats, using individual and pooled animal samples.

One of the primary objectives of toxicogenomics is the development of gene profiles characteristic of discrete toxicities that are potentially measurable on any well-designed and well-annotated microarray platform (Hamadeh et al. 2002; Thomas et al. 2001; Waring et al. 2001). The development of toxicity profiles and specific gene expression signatures requires the use of standardized conditions and the testing of a number of drugs of distinct classes and mechanisms, and therefore they are usually developed on one platform technology. Analytical validation of the veracity of gene changes within a signature is usually performed with an independent gene expression detection technology such as quantitative reverse transcriptase–polymerase chain reaction (qRT–PCR). An additional approach to analytical gene profile validation involves the detection of the identified gene signature on microarray platforms that use different types of probes to capture and detect specific RNA transcripts within samples. Deposition in a public repository of genomic data annotated with toxicity data that was generated by the ILSI working groups is one of the goals of this project (Robinson et al. 2003). The potential to mine gene expression data from diverse sources that will eventually populate public databases will be unrealized if gene profiles of toxicity cannot be extrapolated from one microarray platform to another. It is encouraging that meta-analysis of microarray data collected from different sources and on different microarray formats has shown that cohorts of genes dysregulated in prostate cancer can be consistently identified across data sets (Rhodes et al. 2002). However, more effort must be devoted to evaluating and improving data comparisons across platforms, including verifying probe annotations, gene family and splice variant analysis, and comparability of statistical methods.

In this article, we examine the cross-platform and cross-experiment reproducibility of gene changes induced in rat kidney by cisplatin treatment and identified as statistically significant on a custom cDNA array. Of the studies performed by the HESI Nephrotoxicity Working Group, the most complete data set with the largest number of microarray formats was available for the study involving the antineoplastic agent and regionally specific nephrotoxicant cisplatin. RNA prepared from the kidneys of rats on day 7 after a single injection of 5 mg/kg cisplatin, a dose and time point coincident with proximal tubule necrosis and regeneration, was tested on three different cDNA-spotted arrays (the NIEHS custom rat cDNA array, the Incyte Rat Toxicology GEM, and the PHASE-1 Rat 700 array) and one high-density oligonucleotide array (the Affymetrix RGU34A array).

Materials and Methods

Animal studies. Sprague-Dawley rats Crl:CD(SD)IGSBR VAF/Plus rats were

This article is part of the mini-monograph "Application of Genomics to Mechanism-Based Risk Assessment."

Address correspondence to K.L. Thompson, HFD-910, Division of Applied Pharmacology Research, CDER, U.S. FDA, Life Sciences Building 64, 10903 New Hampshire Ave., Silver Spring, MD 20993 USA. Telephone: (301) 796-0126. Fax: (301) 796-9818. E-mail: Thompsonk@cder.fda.gov

We thank and acknowledge J. Collins, J. Miller, J. Chou, and P. Bushel at the National Center for Toxicogenomics, Research Triangle Park, North Carolina, for the BLAST analysis of cDNAs on the NIEHS platform.

The authors declare they have no competing financial interests.

Received 15 August 2003; accepted 8 January 2004.

purchased from Charles River Laboratories (Raleigh, NC) and treated with cisplatin (CAS no. 1566-27-1; Sigma Chemical Co., St. Louis, MO) or vehicle in two independent studies. A preliminary study was contracted by the National Institute of Environmental Health Sciences (NIEHS) at Integrated Laboratory Systems, Inc. (Research Triangle Park, NC) using a single dose of cisplatin to determine the feasibility of and to direct the design of the overall project. The single time point (on day 7) and dose level (5 mg/kg) serve as a biological replicate for the longest time point of the high-dose group in the multidose study. Serum chemistry parameters were statistically evaluated by a pairwise *t*-test following the conduct of a Bartlett's test for homogeneity of variance. A Dunnett's *t* test was used for those end points for which the Bartlett's test was not significant (p-value > 0.05).

The multidose time course study was performed at Pfizer's Groton laboratories (Groton, CT). Groups of five animals were dosed once by ip injection with 0, 0.3, 1.0, or 5.0 mg/kg cisplatin in saline. Animals were sacrificed at 4, 24, 48, and 144 hr after dosing. Kidneys were snap frozen in liquid nitrogen for RNA isolation. Kidneys were also sampled, fixed in 10% neutral buffered formalin, and blocked in paraffin. Slides from the kidney blocks were processed for hematoxylin and eosin staining.

RNA extraction. Homogenates were prepared from entire kidneys and processed using Qiagen RNeasy Maxi kits (Qiagen, Valencia, CA). Purified total RNA was quantitated by ultraviolet light spectrophotometric determination. Purity and quality were confirmed on gels and an Agilent 2100 bioanalyzer (Agilent Technologies, Palo Alto, CA). For pooled samples, equal amounts of RNA were combined from each animal in the treatment group.

NIEHS custom cDNA array analysis. RNA was labeled, hybridized to microarrays, scanned, and images processed as described (Amin et al. 2004). Briefly, total RNA (35–75 μg) from control animals was labeled with Cyanine 3 (Cy3)-conjugated dUTP and total RNA from treated animals was labeled with Cyanine 5 (Cy5)-conjugated dUTP using SuperScript reverse transcriptase (Invitrogen, Carlsbad, CA). The labeled products from the control and treated animal RNAs were mixed and hybridized to cDNA microarrays in quadruplicate. The cDNA arrays were scanned with an Axon scanner (Axon Instruments, Foster City, CA) and image analysis was conducted using the ArraySuite, version 2.0, extensions of the IPLab image processing software package (Scanalytics, Fairfax, VA). Genes significantly changed at 144 hr by 5 mg/kg cisplatin were determined as outliers in at least three of four fluor flip experiments with a 95% confidence interval. The binomial probability of chance occurrence at the ratio outlier frequency queried was $p < 0.000475$.

Affymetrix RGU34A array analysis. Sample labeling, microarray hybridization, washing, and scanning were performed according to the manufacturer's protocols (Affymetrix, Inc., Santa Clara, CA). Briefly, 10 μg total RNA was reverse transcribed using a T7-$(dT)_{24}$ oligomer and SuperScript. Biotin-labeled cRNA targets were prepared from cDNA using T7 polymerase (Enzo Diagnostics, Inc., Farmingdale, NY), fragmented, and hybridized to the oligonucleotide arrays. After staining with R-phycoerythrin streptavidin (Molecular Probes, Inc., Eugene, OR), arrays were scanned on an Affymetrix GeneChip Scanner 2500 at a photomultiplier tube (PMT) setting of 1,500. The scans from each array were globally scaled by setting the average signal intensity to a target signal of 500. The data were analyzed using Affymetrix Microarray Suite, version 5.0 (MAS 5.0). For samples from individual animals, pair-wise comparisons between each treated animal and each of the five control animals were performed; the signal log ratios were averaged across comparisons; and change calls were based on 60% or greater agreement across the five comparisons.

PHASE-1 Rat 700 microarray analysis. Sample labeling, microarray hybridizations, and washes were performed according to the protocol supplied by the manufacturer (PHASE-1 Molecular Toxicology, Inc., Santa Fe, NM). Twenty μg of total RNA was primed with oligo dT and labeled by direct incorporation of either Cy3-dCTP or Cy5-dCTP during reverse transcription. cDNA from labeled treated and control samples was purified on Wizard Series 9600 DNA purification resin (Promega Corp., Madison, WI), combined, and hybridized to the same microarray. Each microarray was scanned using a ScanArray 4000 microarray scanner (PerkinElmer, Inc., Wellesley, MA). PMT values were adjusted to balance total fluorescence signal between the two channels. QuantArray software, version 2.0 (PerkinElmer, Inc.), was used to quantitate the bound signal for each spot on the microarray from the ScanArray output TIFF file. Local background fluorescence was subtracted and the resulting data Lowess normalized. The calculated $\log_2$ ratios were averaged from quadruplicate spots deposited for each queried element on each array.

Incyte Rat Toxicology GEM analysis. Transcription profiling using Incyte cDNA gene expression microarrays (GEMs; Incyte Corp., Palo Alto, CA) was performed essentially as previously described (Schena et al. 1996). Briefly, poly(A)$^+$ RNA was isolated using MicroPoly(A)Pure kits (Ambion, Inc., Austin, TX). Aliquots of 200 ng mRNA were used to generate Cy3- and Cy5-labeled cDNA probes using GEMBrite probe labeling kits (Incyte Corp.). Individual treated and control sample probes were labeled with Cy5, while a pool of equal proportions of the control group mRNAs was used to generate Cy3-labeled cDNA probes. Probe hybridization using Rat Toxicology GEMs was performed at Incyte Corp. as previously described (Schena et al. 1996). Each slide was analyzed using an Incyte internal program. Microarray hybridizations were evaluated with respect to five quality control parameters, background correction, log balance coefficient, absolute average signal, differential expression ratio, and differential percentage. Signal intensities in the Cy3 and Cy5 channels were normalized using total average signal intensity in both channels to generate a balance coefficient. Elements with signal intensities P1 + P2 < 500, signal to background values $P1_{STB}$ + $P2_{STB}$ < 10, or area of coverage < 40%, were recorded as absent values and were omitted from further consideration.

The complete data set is currently being submitted to ArrayExpress (European Bioinformatics Institute, Hinxton, UK; http://www.ebi.ac.uk/arrayexpress) and will be available for public download by the second quarter of 2004. Accession numbers referencing this data set will be available on the HESI web site (http://hesi.ilsi.org/index.cfm?pubentityid=120).

Real-time qPCR. Primers for heme oxygenase-1 (*Hmox1*), clusterin (*Clu*), osteopontin (*Spp1*), and solute carrier family 15, member 2 (*Slc15a2*) were designed using Primer Express software (Applied Biosystems, Foster City, CA) and custom made (Research Genetics, Huntsville, AL). The primer sequences (5′ to 3′) were as follows: *Hmox1* forward CGTGCTCG CATGAACACTCT, *Hmox1* reverse CGGTCTTAGCCTCTTCTGT, *Clu* forward TGTGGACTGTTCGACCAACAA, *Clu* reverse GAGGGAATGAAGCAGCT CGTT, *Spp1* forward CCAGCCAAGGA CCAACTACAA, *Spp1* reverse GCCAAA CTCAGCCACTTTCAC, *Slc15a2* forward GCAATGGGAAGCAAGATGTACAG, and *Slc15a2* reverse GTGTTGTCGCT TCGGAAGGT. Real-time PCR was performed using the ABI Prism 7700 Sequence Detection System (Applied Biosystems) according to the manufacturer's instructions. The SYBR Green I labeling kit (Applied

Biosystems) was used to detect double-stranded DNA generated during PCR amplification according to manufacturer's instructions. Reverse transcription and PCR reactions were performed simultaneously in a 50-µL reaction containing 4 mM $MgCl_2$, 0.8 mM of each dNTP, 100 ng total RNA, 0.4 µM reverse primer and 0.4 µM forward primer, 0.4 U/µL Rnasin, 0.025 U/µL AmpliTaq Gold DNA polymerase (Roche, Basel, Switzerland) and 0.25 U/µL MuIV reverse transcriptase (Roche). Amplification reactions were carried out using the following temperature profile: 48°C, 10 min; 95°C, 15 sec; 60°C, 1 min) for 40 cycles. Fluorescence emission was detected for each PCR cycle and the threshold cycle (C_T) values were determined. The C_T value was defined as the actual PCR cycle when the fluorescence signal increased above the background threshold. Induction or repression of a gene in a treated sample relative to control was calculated as follows: fold increase/decrease = $2 - [C_{T\ (exposed)} - C_{T\ (control)}]$. Values were reported as an average of triplicate analyses normalized to glyceraldehyde phosphate dehydrogenase RNA levels.

Results

Cross-platform comparisons. As part of the studies conducted by the HESI Nephrotoxicity Working Group, Sprague-Dawley rats were injected with a single dose of cisplatin (0, 0.3, 1, or 5 mg/kg) and sacrificed after 4, 24, 48, or 144 hr. RNA prepared from the kidneys of these animals was analyzed on several different microarray platforms at different sites. The platforms involved in this study include the NIEHS custom array spotted with 7,000 cDNAs and expressed sequence tags (ESTs), the Affymetrix RGU34A array that contains 8,800 *in situ* synthesized oligonucleotide probe sets, the Incyte Rat Toxicology GEM that has 9,984 spotted cDNAs and ESTs, and the PHASE-1 Rat Expression array with 700 spotted cDNAs and ESTs. The three cDNA platforms were all constructed using PCR products of 500–2,000 nucleotides in length amplified from libraries of rat cDNA clones. The RGU34A array was constructed using 16 pairs of perfect match and single base-pair (bp) mismatched oligonucleotides, each 25 bp in length, to interrogate each transcript.

A set of 93 genes was identified that was significantly altered in the kidney after exposure to a nephrotoxic dose of cisplatin (5 mg/kg for 144 hr) on the NIEHS array using four independent dye-swap measurements of RNA pooled from individual animals (Amin et al. 2004). The pooled high-dose cisplatin sample was analyzed at different sites on the Affymetrix, Incyte, and PHASE-1 platforms only in single chip determinations. However, biological replicates for this data point were run on all of these three different platforms. Because biological variability is thought to be greater than technical variability for well-performed microarray experiments (Yang and Speed 2002), the data from the biological replicates were used for the cross-platform comparison. RNAs from the five rats in the 144-hr high-dose cisplatin group and the five rats in the 144-hr control group were run on individual RGU34A arrays in singlicate. On the Incyte and PHASE-1 platforms, individual treated animal samples were run in singlicate on separate arrays, paired with pooled control RNA labeled with a different fluor. It was readily apparent from tallying the number of gene changes above 2-fold on any of the three platforms that the individual animals did not respond homogeneously to cisplatin. These observations were also reflected in the traditional toxicologic end points of histopathology and serum markers indicative of renal injury. Histopathologic diagnosis revealed that three animals (nos. 76, 77, and 78) had similar moderate levels of single-cell necrosis and tubular regeneration in the outer medulla of the kidney, one animal (no. 79) had a mild degree of injury, and one animal (no. 80) had no visible kidney damage (Table 1). The blood urea nitrogen (BUN) and serum creatinine levels were significantly elevated in animal nos. 76, 77, and 78 but were at control levels in animal nos. 79 and 80.

Cross-platform comparisons can be ascertained best when probe sequences on the platforms used are fully and correctly annotated. The most in-depth comparison of probes between platforms was performed for the NIEHS and Affymetrix RGU34A arrays, for which the most complete publicly available sequence information was available. The NIEHS array is printed with PCR-amplified sequences from sequence-verified Research Genetics rat cDNA library clones (Amin et al. 2004). The Affymetrix rat genome U34A (RGU34A) oligonucleotide array was designed to interrogate 7,000 full-length sequences and about 1,000 EST clusters from the UniGene (http://www.ncbi.nih.gov/UniGene/) database (Build 34; Affymetrix Inc. 2001). The probes of interest were first cross-referenced between the two platforms by UniGene identifier. Where multiple probe sets on the RGU34A array had the same UniGene identifier as a probe of interest on the NIEHS array, the Probe Match tool available on NetAffx (http://www.affymetrix.com/analysis/inde.affx; Liu et al. 2003) was used to determine which RGU34A probe set had the maximal sequence overlap with the spotted NIEHS sequence. Although probe sequence overlap was not necessary for good correlation of signal between platforms, it was used here to develop the closest map between probe sequences on the two platforms. Additionally, the Probe Match tool was used to identify Affymetrix probe sets that best corresponded to NIEHS probes that had not been assigned UniGene identifiers or had UniGene identifiers that did not match the annotations available for genes on the RGU34A array. Finally, if there was poor agreement in performance for probes matched across platforms by UniGene identifier, the Probe Match tool was used to query for better sequence matches between NIEHS probes and RGU34A probe sets. Probes of interest on the NIEHS array were matched to probes on the Incyte Rat Toxicology GEM and PHASE-1 Rat 700 arrays by UniGene identifier (Mattes 2004).

Forty-eight of the 93 differentially expressed genes identified in the NIEHS data set could be matched to probe sets on the Affymetrix RGU34A array by UniGene identifier or sequence match. An additional 34 genes in the NIEHS data set could be matched to probe sets on the RGU34B and C arrays, which were not used in this study. For 11 of the 93 genes, either no corresponding RGU34 probe set was found or the sequence spotted on the NIEHS array could not be confirmed by sequence comparison against the National Center for Biotechnology Information (NCBI) nucleotide database using Basic Local Alignment Search Tool (BLAST) analysis (Altschul et al. 1997). Using the annotation provided by the manufacturer supplemented by BLAST analysis for a subset of probes, 33 of the 93 genes in the NIEHS data set were matched to probes on the Incyte Rat Toxicology GEM.

Table 1. Variation in individual animal histopathology diagnosis and serum markers of nephrotoxicity.

	Individual animal no.				
	76	77	78	79	80
Single cell necrosis, outer medulla	Moderate	Moderate	Moderate	Mild	None
Tubular degeneration	Mild	Mild	Mild	Mild	None
Tubular regeneration	Moderate	Moderate	Moderate	Mild	None
BUN (mg/dL)[a]	87	100	106	11	13
Serum creatinine (mg/dL)[b]	1.4	2.4	2.3	0.4	0.3

[a]Controls (all time points): 13.05 ± 2.4 (average ± SD, $n = 20$). [b]Controls (all time points): 0.28 ± 0.04 (average ± SD, $n = 20$).

Sixteen of the 93 could be matched to probes on the PHASE-1 rat subarray.

Gene expression data from each of the three platforms for the three animals with similar biological response to cisplatin were averaged and compared with the response determined using technical replicates of pooled animal data on the NIEHS platform (Tables 2 and 3). Among the 48 genes that could be linked between the NIEHS and the RGU34A platform, a good agreement in the direction, statistical significance, and magnitude of change was observed for the majority of the genes (88%) in the data set. Ninety-four percent of the 33 genes in the NIEHS data set identified as present on the Incyte Rat Toxicology GEM were also consistently

Table 2. Genes associated with renal injury by cisplatin on multiple platforms.

UniGene ID	NIEHS ID[a]	Gene symbol[b]	Gene name[b]	NIEHS[c]	Affymetrix[d]	Call[e]	Incyte[d]	PHASE-1[d]
Rn.1780	AA818413	*Clu*	Clusterin	2.895	3.289	I	3.360	2.865
Rn.8871	AA964431	*Spp1*	Secreted phosphoprotein 1	2.064	2.173	I	2.599	1.312
Rn.2710	AA859385	*Vim*	Vimentin	1.642	2.278	I	1.498	Absent[f]
Rn.90546	AA964578	*Anxa2*	Calpactin I heavy chain	1.328	2.167	I	2.140	2.005
Rn.2458	AA858888	*Tubb5*	Tubulin, beta 5	1.263	1.490	I	0.829	Absent
Rn.87063	AA955668	*Gstp2*	Glutathione *S*-transferase, pi 2	1.239	2.206	I	1.549	1.740
Rn.98846	AA925421	*Fga*	Fibrinogen, alpha polypeptide	1.227	1.914	I	1.249	1.076
Rn.16933	NM053380	*Slc34a2*	Solute carrier family 34, member 2	1.118	Absent		0.607	Absent
Rn.5983	AA924288	*Tmsb10*	Thymosin beta 10	1.091	1.822	I	0.894	0.926
Rn.11303	NM130741	*Lcn2*	Lipocalin 2	1.057	3.943	I	Absent	Absent
Rn.5834	AA925280	*Ccng1*	Cyclin G1	0.978	2.605	I	1.355	1.505
Rn.24945	M31109	*Udpgt*	UDP-Glucuronsyltransferase 2B3	0.926	0.721	I	Absent	Absent
	L12458	*Lyz*	Lysozyme	0.918	0.933	I	Absent	Absent
Rn.2521	NM031533	*Ugt2b*	UDP-Glucuronosyltransferase 2 family, polypeptide B	0.895	0.503	I	Absent	Absent
Rn.2605	AA819102	*Tmsb4x*	Thymosin beta-4	0.895	0.885	I	1.191	Absent
Rn.4083	AA900235	*S100a10*	S-100 related protein, clone 42c	0.848	1.700	I	1.140	Absent
Rn.1119	NM022509	*Smn*	Survival motor neuron	0.782	1.019	I	Absent	Absent
Rn.2379	AA874853	*Mgp*	Matrix Gla protein	0.766	0.799	I	0.711	Absent
Rn.625	AA998734	*Gstm2*	Glutathione *S*-transferase, mu 2	0.740	1.259	I	1.144	0.039
Rn.764	AA859797	*Lgals3*	Lectin, galactose binding, soluble 3	0.678	2.060	I	1.305	0.328
Rn.2953	AA924727	*Col1a1*	Collagen type 1, alpha 1	0.614	0.569	I	0.598	Absent
Rn.3160	AA874884	*Hmox1*	Heme oxygenase 1	0.566	1.092	I	Absent	0.455
Rn.3545	AI137902	*Tgfb1i4*	Transforming growth factor beta 1 induced transcript 4	0.526	0.671	I	0.737	Absent
Rn.10992	AA964628	*G6pc*	Glucose-6-phosphatase	–0.578	–1.370	D	–1.270	Absent
Rn.26060	AA997886	*Cyp2d18*	Cytochrome P450 2D18	–0.578	–0.421	NC	–0.692	–1.393
Rn.2178	AA819745	***Ghr***	Growth hormone receptor	–0.599	–1.340	D	–0.857	Absent
Rn.2589	AA818579	*Cdo1*	Cysteine dioxygenase 1	–0.644	–1.627	D	–0.939	Absent
Rn.103392	AA924904		EST	–0.644	–1.373	D	Absent	Absent
Rn.11132	AA998088	*Anpep*	Alanyl aminopeptidase	–0.667	–1.798	D	Absent	Absent
Rn.1437	AA818706	*Gc*	Group-specific component	–0.713	–2.223	D	Absent	Absent
Rn.10417	AI029336	*Hao3*	Hydroxyacid oxidase 3	–0.737	–0.501	D	–1.033	Absent
Rn.1874	AA818440	*AGT2*	Beta-alanine-pyruvate aminotransferase	–0.737	–0.779	D	–0.867	Absent
Rn.33890	AA819309	*Gamt*	Guanidinoacetate methyltransferase	–0.737	–1.265	D	Absent	Absent
Rn.89268	AA818855	*Slc15a2*	Solute carrier family 15, member 2	–0.737	–2.527	D	–1.165	Absent
Rn.26369	AA819611	*Igfbp3*	Insulin-like growth factor binding protein 3	–0.761	–1.409	D	–0.758	–0.905
Rn.54567	AA818636	*Siat1*	Sialyltransferase 1	–0.761	–1.429	D	–1.061	Absent
Rn.9230	AA957263		EST, similar to connexin protein Cx26	–0.761	–1.203	D	Absent	Absent
Rn.11331	AA925291	*Ngfg*	Nerve grown factor, gamma	–0.837	–1.626	D	–1.566	Absent
Rn.874	AA860013	*Odc1*	Ornithine decarboxylase 1	–0.837	–1.383	D	Absent	–1.176
Rn.11143	AI070884	*Kap*	Kidney androgen-regulated protein	–0.889	–1.806	D	–2.100	Absent
Rn.1430	AA818680	*Oat*	Ornithine aminotransferase	–0.889	–1.930	D	–1.177	–0.326
Rn.108214	AA858962	*Rbp4*	Retinol binding protein 4	–0.889	–2.220	D	Absent	–1.020
Rn.11133	AA998607	*Kat2*	Kynurenine aminotransferase 2	–0.971	–1.830	D	–2.043	Absent
Rn.6075	AA901327	*Egf*	Epidermal growth factor	–1.218	–2.581	D	–1.905	–1.905

Abbreviations: D, decrease; I, increase; NC, no change.
Values are fold inductions shown as $\log_2$ ratios. [a]Underlined NIEHS ID is best BLAST match; nonunderlined NIEHS ID is spotted sequence. [b]Annotated from the UniGene database (http://www.ncbi.nih.gov/UniGene/). [c]Average of four technical replicates of pooled RNA. [d]Average of three biological replicates of individual animal RNAs. [e]Call by MAS 5.0 analysis. [f]"Absent" indicates that the probe sequence is not on the microarray.

Table 3. Genes associated with renal injury by cisplatin that were not confirmed on multiple platforms.

UniGene ID	NIEHS ID[a]	Gene symbol[b]	Gene name[b]	NIEHS[c]	Affymetrix[d]	Call[e]	Incyte[d]	PHASE-1[d]
Rn.3603	AA900551	*Ephx1*	Epoxide hydrolase 1	0.687	0.506	NC	0.479	Absent[f]
Rn.8707	AA955199	*Gdf10*	Prepro bone inducing protein	0.642	–0.198	NC	Absent	Absent
	AA899443		EST, similar to mitochondrial cytochrome B	–0.667	–0.291	NC	Absent	Absent
Rn.19270	AI059765	*Anxa4*	ZAP 36/annexin IV	–0.713	0.622	I	Absent	Absent
Rn.10958	AI059692	*Eef2k*	Eukaryotic elongation factor 2 kinase	–0.761	–0.002	NC	0.163	Absent

Abbreviations: D, decrease; I, increase; NC, no change.
Values are fold inductions shown as $\log_2$ ratios. [a]Underlined NIEHS ID is best BLAST match; nonunderlined NIEHS ID is spotted sequence. [b]Annotated from the UniGene database. [c]Average of four technical replicates of pooled RNA. [d]Average of three biological replicates of individual animal RNAs. [e]Call by MAS 5.0 analysis. [f]"Absent" indicates that the probe sequence is not on the microarray.

changed > 1.5-fold among the biological replicates assayed on that platform. Seventy-five percent of the 16 genes in the NIEHS data set that matched to the PHASE-1 Rat 700 array showed a reproducible change > 1.5-fold among the biological replicates. Five of the genes that were significantly changed on the NIEHS platform in the kidney samples lacked concordance of response with the matched probes on other platforms (Table 3). Correct annotation of the five probe sequences on the NIEHS platform was reconfirmed by BLAST analysis. For 3 of the 5 probes, there was complete sequence overlap between the 16 probe pairs in the most homologous Affymetrix probeset identified and the cDNA sequence. For 2 probes, there was partial overlap between the cDNA and Affymetrix probe pair sequences. Discordant outcomes between platforms may result from cross-hybridization on the cDNA platforms to nontarget sequences, differential detection of alternate splice variants by probes from the same UniGene cluster, misannotation of probes on arrays, mistakes in probe placement on arrays, or large differences in hybridization affinity between probe sequences from the same UniGene cluster.

Table 4. Correlation between cisplatin-altered gene changes in pooled animal samples on the NIEHS platform and individual animal samples for probes common to other individual platforms.

Platforms compared with NIEHS array	Individual animal number					Pool	No. of genes[a]
	76	77	78	79	80		
Affymetrix RGU34A	0.896[b]	0.881	0.874	0.655	–0.141	0.781	48
Incyte Rat Toxicology GEM	0.942	0.940	0.941	0.830	–0.378	NA	33
PHASE-1 Rat Array	0.948	0.930	0.912	0.803	–0.333	0.972	16

NA, not available. [a]Number of genes in comparison. [b]All values are Pearson correlation coefficients.

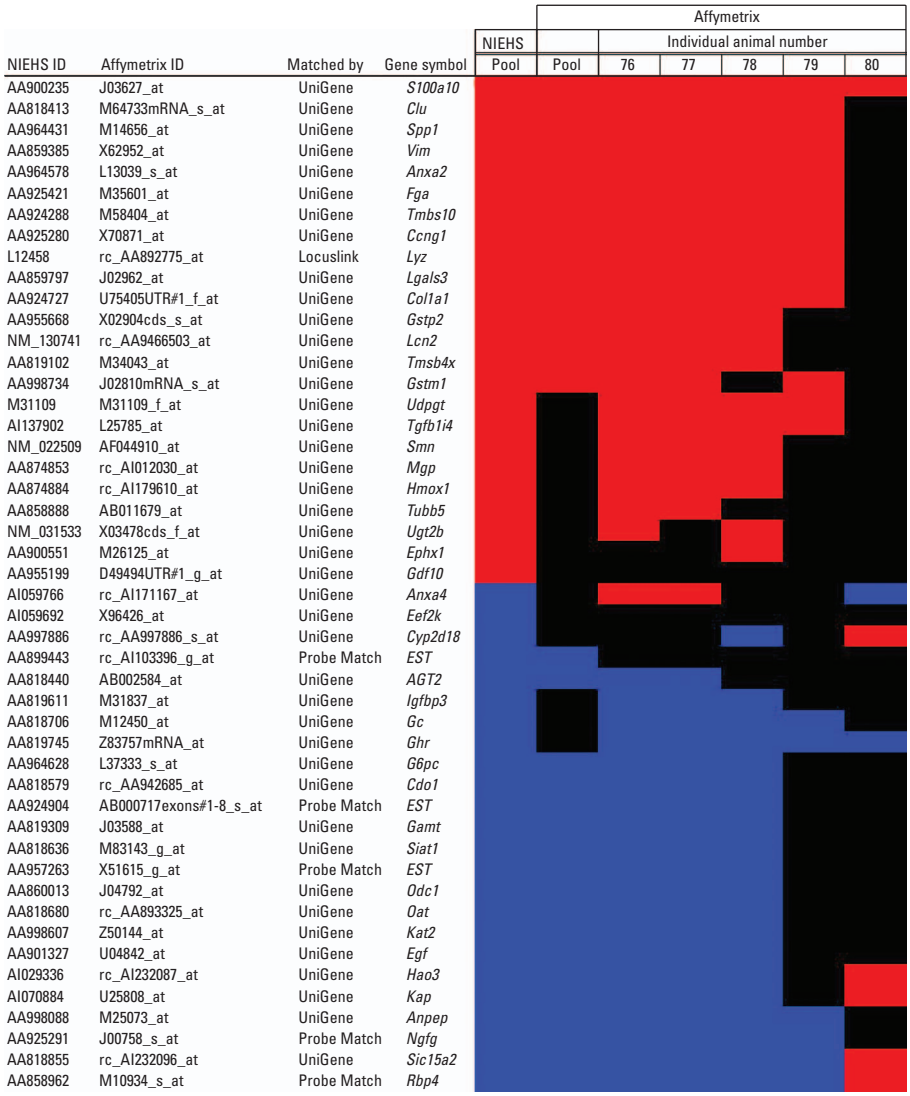

Figure 1. Comparison of gene expression results for the intersection of cisplatin-altered genes on the NIEHS cDNA and Affymetrix RGU34A platforms. The direction of the statistically significant change in pooled or individual animal analyses is indicated by color. Blue represents decreased expression; red represents increased expression; black indicates not changed, as determined by MAS 5.0 analysis on Affymetrix arrays.

A strong concordance was observed between the degree of renal injury induced by cisplatin in individual animals, as determined by traditional toxicological end points and the strength of correlation of the corresponding gene expression results between platforms. RNA assayed on the RGU34A platform from the three animals (nos. 76–78) that incurred a similar, moderate level of renal injury from cisplatin showed the highest concordance of response with the results on the NIEHS array determined using RNA pooled from the five individual animals in this group (Table 4). Of the 48 genes in the data set cross-matched between the NIEHS and RGU34A platforms, 41–42 genes from each of animals nos. 76, 77, and 78 were assigned change calls using MAS 5.0 analysis of RGU34A arrays that agreed with the results from quadruplicate measurements of pooled RNA on the NIEHS array (Figure 1). The Pearson correlation coefficients between the individual animal data determined on oligonucleotide arrays and the pooled data determined on cDNA arrays showed relatively good correlation and close agreement between biological replicates (Table 4). Kidney RNA from animal no. 79, which had a milder response to cisplatin, displayed fewer significant changes among the genes indicative of proximal tubular injury. The Pearson correlation coefficient for the comparison between the $\log_2$ ratio data for animal no. 79 and the pool assayed on the NIEHS platform was lower (0.655) and only 20 of the 48 genes had change calls in agreement in this comparison. In the kidney RNA from animal no. 80, which showed no histopathology response to cisplatin, only 3 of the 48 genes in the data set altered after cisplatin treatment were significantly changed from control levels and the Pearson correlation coefficient was –0.141 between the $\log_2$ ratio data for animal no. 80 and the NIEHS pool. Similar correlation between gene expression and degree of renal injury for the five individual animals treated with 5 mg/kg cisplatin for 144 hr were seen using the Incyte Rat Toxicology GEM and PHASE-1 Rat Array (Table 4).

Individual versus pooled animal results. The presence of one apparent nonresponder and one weakly responding animal among

five animals in a treatment group might be expected to dilute the signal of genes changed by renal injury in pooled samples. This result was observed on the RGU34A array (Figure 1, Table 4). Specifically, a lower Pearson correlation coefficient and fewer MAS 5.0 change calls were seen with pooled RNA than with samples from the individual animals displaying the full response to cisplatin (32 vs. ~ 41). Although a similar loss of signal was not seen with the PHASE-1 platform, this could be due to greater technical variation. The individual animal samples were hybridized on a different day than the pooled sample, and it has been observed that dye effects seen with this platform are most consistent within hybridization batch (Rosenzweig et al. 2004). Although the cisplatin-induced data set was determined on the NIEHS platform using pooled animal samples, the statistical power added by the use of four replicates might compensate for some loss of signal due to pooling. However, if a uniform response is not seen among animals within an individual treatment group as was observed in this study, pooling may attenuate measurable gene expression changes and conceal small but biologically relevant changes.

Verification by qRT–PCR of differentially expressed genes. The changes in expression assayed on multiple microarray formats in kidneys with proximal tubule damage were further verified for 4 genes by qRT–PCR assay (Table 5). In the pooled animal sample, heme oxygenase-1 (*Hmox1*) was induced almost 8-fold ($\log_2$ ratio of 3) in the qRT–PCR assay but 2-fold or less on the oligonucleotide- and cDNA-based platforms. In contrast, *Clu* was induced about 3-fold in the qRT–PCR assay and 6- to 10-fold on microarrays. *Spp1* and *Slc15a2* genes were induced or repressed to similar extents in the qRT–PCR assay and on all platforms assayed in this study where present.

Repeatability of* in vivo *studies. The cross-study applicability of the set of gene markers of cisplatin-induced injury was examined using data from the preliminary study conducted by the Nephrotoxicity Working Group at an independent site from the multidose study, including RNA preparation. For the preliminary study, six animals per treatment group received a single injection of saline or 5 mg/kg cisplatin and were sacrificed 144 hr later. All animals in the group displayed significantly lower body weights and liver weights at sacrifice, but only slight elevations were observed in BUN (28 mg/dL ± 7 mg/dL, mean ± SD, in treated vs. 17 mg/dL ± 1 mg/dL in controls; Cochran-Cox *t*-test pairwise *p*-value of 0.01) and serum creatinine (0.6 mg/dL ± 0.06 mg/dL in treated vs. 0.25 mg/dL ± 0.08 mg/dL in controls; Dunnett's test pairwise *p*-value < 0.05). The RNA from the animals was pooled within the treated and control groups and tested on PHASE-1 arrays on six replicate arrays with flipped fluor labeling. Among the 16 genes in the NIEHS set of cisplatin-altered genes that match to probes on the PHASE-1 platform, a similar trend in gene expression change was observed between replicate experiments on the same array and across arrays (Figure 2), even though serum markers of renal dysfunction were altered to different extents in the two studies. One gene (*Gstm2*) that showed a significant change only on the NIEHS platform and not on the PHASE-1 platform in either replicate study was discovered on closer inspection to have been originally misannotated on the NIEHS array. Recent BLAST analysis of the spotted NIEHS sequence at this location revealed that the closest sequence match for this probe was with *Gstm1*, not *Gstm2*. Probes correctly annnotated and matched between the NIEHS and Phase-1 arrays might be anticipated to perform similarly with the same samples if differences in probe labeling, hybridization, scanning, and data processing do not mask the true gene expression changes within the samples because both arrays were constructed from the same library of rat cDNA clones.

Discussion

In this study we examined the ability of different microarray formats to identify gene expression changes in kidneys of rats treated with cisplatin. A set of differentially expressed genes that were significantly changed was developed from four replicate determinations on a custom cDNA array. We found that a significant subset of the genes in this set could be identified as differentially expressed on each of the other platforms, transcending specific target capture technologies and providing additional validation to their inclusion in the profile. These data also suggest that the same biology can be captured and reported, independent of gene coverage, by different array formats when gene profiles are developed that are linked to defined toxicities.

Many of the 43 genes in the platform-independent gene profile have known

Table 5. Verification by qRT–PCR of cisplatin-altered gene changes measured on multiple microarray formats.

Gene symbol	qRT–PCR	NIEHS	Affymetrix	Incyte[a]	PHASE-1
Hmox1	2.74	0.56	1.09	Absent[b]	0.65
Clu	1.72	2.90	2.59	2.43	3.26
Spp1	2.38	2.06	2.11	1.95	1.84
Slc15a2	–1.00	–0.74	–1.14	–0.83	Absent

Values are $\log_2$ fold inductions for pooled animal samples. [a]Average of five biological replicates (not determined for pooled animal samples). [b]"Absent" indicates that the probe sequence is not on the microarray.

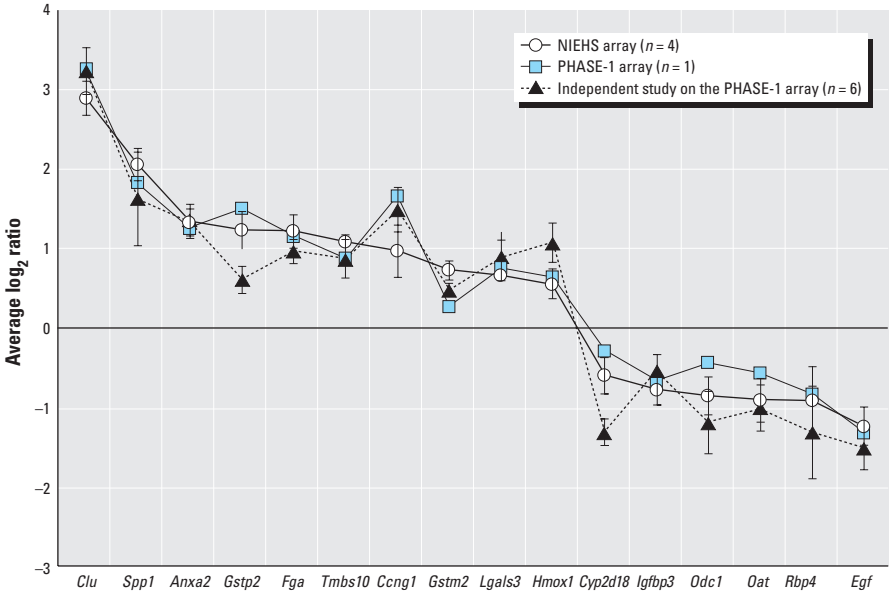

Figure 2. Comparison of average $\log_2$ ratios between independent *in vivo* experiments. Average $\log_2$ ratios for the 16 genes significantly altered in rat kidney by 5 mg/kg cisplatin on the NIEHS custom cDNA array and cross-matched to the PHASE-1 Rat Toxicology array are shown for pooled animal samples from the same study on the NIEHS array and PHASE-1 array and for an independent study on the PHASE-1 array. Genes are designated by locus symbol.

structural or functional ontologies that are consistent with the biology associated with cisplatin nephrotoxicity. Genes associated with damage to proximal tubules, tissue remodeling, and regeneration are represented in this set. Genes involved in cytoskeletal structure and function (*Vim, Tubb5, Tmsb10, Tmsb4x, Anxa2*), in cell adhesion (*Spp1, Col1a1, Clu, Lgals3*), and detoxification enzymes (*Gstm2, Gstp2*) are upregulated following cisplatin-induced injury. Genes downregulated by cisplatin include those that localize to the proximal tubules (*Odc1, Oat, G6pc, Kap*) and those that encode growth factors or their binding proteins (*Egf, Ngfg, Igfbp3, Ghr*).

Further global analyses of gene expression changes associated with different types of nephrotoxicity will be necessary to define which genes in this set are hallmarks of specific mechanisms of injury and which are more general markers of renal damage, immune cell invasion, or tissue remodeling. Subsets of the genes identified as platform-independent markers of cisplatin-induced renal damage in this study have been observed in other global analyses of differentially expressed genes after renal damage. Significant changes in the expression of *Clu, Tmsb10, Ccng1, Igfbp3*, and *Egf* were observed in kidney RNA from male Sprague-Dawley rats that received daily injections of 1 mg/kg cisplatin for 7 days using PHASE-1 Rat 250 microarrays (Huang et al. 2001). Upregulation of *Tubb5, Col1a1, Gst, Lgals3,* and *Anxa2* and downregulation of *Igfbp3* and *Egf* were also observed in ischemia-induced acute renal failure in mice on Affymetrix Murine U74A arrays (Yoshida et al. 2002). Calcium oxalate nephrolithiasis induced *Tmsb4x, Col1a1, Spp1*, and *Tgfb1i4* expression in rat kidney analyzed on a custom cDNA microarray (Katsuma et al. 2002).

Analysis of the same sample on different microarray formats can provide an alternate approach for confirming gene expression changes. Discordant results between platforms may identify examples of incorrect annotation or cross-hybridization to non-target gene sequences. Other potential sources of variability such as differences in RNA extraction or other processing methods were not addressed in this study. Earlier cross-platform studies that compared the correlation between all statistically changed genes that can be mapped between different platforms have shown either poor correlation or disparities in sensitivity between platforms (Kuo et al. 2002; Li et al. 2002). The higher degree of correlation between platforms seen in this study may be due to the use of a data set composed of statistically significant gene changes from individual animal replicates that were phenotypically anchored to histopathology and to the use of sequence alignment to match probes between different platforms. As more experience is gained with multiplatform analysis, probes will be identified that are able to detect particular gene transcripts with higher fidelity, enhancing our capability for optimizing measurements of gene expression changes on a genome scale. The current state of knowledge about the rat transcriptome, gene families, and splice variants hinders complete and accurate cross-annotation of the targets detected by probes on different platforms. Continued probing of sequence validation based on cross-platform analyses should help improve the overall quality of microarray gene expression data.

References

Affymetrix Inc. 2001. GeneChip Rat Genome U34 Set Datasheet. Santa Clara, CA:Affymetrix Inc. Available: http://www.affymetrix.com/support/technical/datasheets/rgu34_datasheet.pdf [accessed 26 August 2003].

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402.

Amin RP, Vickers AE, Sistare F, Thompson KL, Roman RJ, Lawton M, et al. 2004. Identification of putative gene-based markers of renal toxicity. Environ Health Perspect 12:465–479.

Hamadeh HK, Bushel PR, Jayadev S, Martin K, DiSorbo O, Sieber S, et al. 2002. Gene expression analysis reveals chemical-specific profiles. Toxicol Sci 67:219–231.

Huang Q, Dunn RT, Jayadev, S, DiSorbo O, Pack FD, Farr SB, et al. 2001. Assessment of cisplatin-induced nephrotoxicity by microarray technology. Toxicol Sci 63:196–207.

Katsuma S, Shijima S, Kirasawa A, Takagaki K, Kaminishi Y, Koba M, et al. 2002. Global analysis of differentially expressed genes during progression of calcium oxalate nephrolithiasis. Biochem Biophys Res Commun 296:544–552.

Kramer JA, Pettit SD, Amin RP, Bertram TA, Car B, Cunningham M, et al. 2004. Overview of the application of transcription profiling using selected nephrotoxicants for toxicology assessment. Environ Health Perspect 12:460–464.

Kuo WP, Jenssen T-K, Butte AJ, Ohno-Machado L, Kohana IS. 2002. Analysis of matched mRNA measurements from two different microarray technologies. Bioinformatics 18:405–412.

Li J, Pankratz M, Johnson JA. 2002. Differential gene expression patterns revealed by oligonucleotide versus long cDNA arrays. Toxicol Sci 69:383–390.

Liu G, Loraine AE, Shigeta R, Cline M, Cheng J, Valmeekam V, et al. 2003. NetAffx: Affymetrix probesets and annotations. Nucleic Acids Res 31:82–86.

Mattes WB. 2004. Annotation and cross-indexing of array elements on multiple platforms. Environ Health Perspect 12:506–510.

Rhodes DR, Barrette TR, Rubin MA, Ghosh D, Chinnaiyan AM. 2002. Meta-analysis of microarrays: interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. Cancer Res 62:4427–4433.

Robinson DE, Pettit SD, Morgan DG. 2003. Use of genomics in mechanisms based risk assessment. In: Toxicogenomics (Inoue T, Pennie WE, eds). Tokyo:Springer-Verlag, 194–203.

Rosenzweig BA, Pine PS, Domon OE, Morris SM, Sistare FD. 2004. Dye-bias correction in dual-labeled cDNA microarray gene expression measurements. Environ Health Perspect 12:480–487.

Schena M, Shalon D, Heller R, Chai A, Brown PO, Davis RW. 1996. Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. Proc Natl Acad Sci USA 93:10614–10619.

Thomas RS, Rank DR, Penn SG, Zastrow GM, Hayes KR, Pande K, et al. 2001. Identification of toxicologically predictive gene sets using cDNA microarrays. Mol Pharmacol 60:1189–1194.

Waring JF, Jolly RA, Ciurlionis R, Lum PY, Praestgaard JT, Morfitt DC, et al. 2001. Clustering of hepatotoxins based on mechanism of toxicity using gene expression profiles. Toxicol Appl Pharmacol 175:28–42.

Yang YH, Speed T. 2002. Design issues for cDNA microarray experiments. Nat Rev Genet 3:579–588.

Yoshida T, Tan S-S, Hsiao L-L, Jensen RV, Ingelfinger JR, Gullans SR. 2002. Global analysis of gene expression in renal ischemia-reperfusion in the mouse. Biochem Biophys Res Commun 291:787–794.